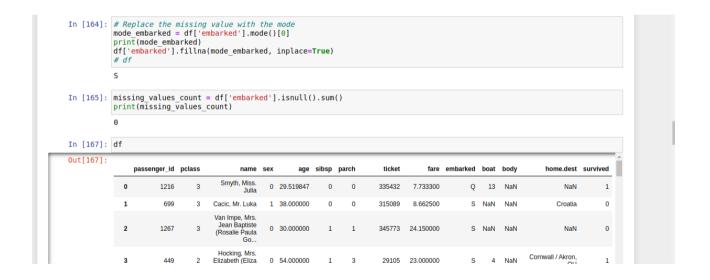
Name: Ali Bilal

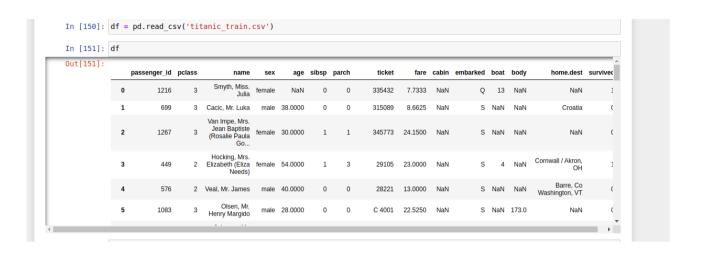
Roll No: P20-0077

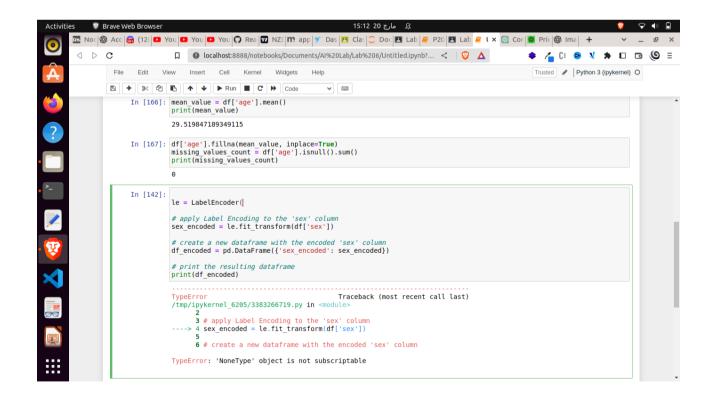
**Section: 6C** 

## **LAB TASK 6**

```
In [188]: import pandas as pd
from sklearn.preprocessing import LabelEncoder
from sklearn.ensemble import RandomForestClassifier
from sklearn.model_selection import train_test_split
from sklearn.neural_network import MLPClassifier
from sklearn.metrics import recall_score, f1_score
```







```
In [152]: # df = df['age']
missing_values_count = df['age'].isnull().sum()
print(missing_values_count)

174

In [153]: mean_value = df['age'].mean()
print(mean_value)
29.519847189349115

In [154]: df['age'].fillna(mean_value, inplace=True)
missing_values_count = df['fare'].isnull().sum()
print(missing_values_count)

1

In [155]: le = LabelEncoder()
# apply_Label_Encoding_to_the_'color'_column_df['sex'] = le.fit_transform(df['sex'])
# print_the_resulting_dataframe
print(df)
```

```
In [171]: df = df.drop('name', axis=1)
In [172]: df = df.drop('boat', axis=1)
In [173]: df
Out[173]:
            passenger_id pclass sex
                               age sibsp parch
                                                      ticket
                                                               fare embarked survived
         0 1216 3 0 29.519847 0 0
                                                  335432 7.733300
                                                                       Q
                                                                             1
                  699
                       3 1 38,000000
                                          0
          1
                                       0
                                                     315089 8.662500
                                                                       S
                                                                              0
         2 1267 3 0 30.000000 1 1
                                                     345773 24.150000
                 449
                                                     29105 23.000000
                      2 0 54.000000
          4 576 2 1 40.00000 0 0
                                                    28221 13.000000 S 0
          5
                 1083
                      3 1 28.000000
                                      0 0
                                                     C 4001 22.525000
                                                                       S
                                                                             0
          6 898 3 1 19.000000 0 0
                                                    LINE 0.000000
                                                                     S 0
                  560
                        2 0 30.000000
                                           0
                                                     250648 13.000000
                                      0
                                                                       S
                                                                              1
                 1079 3 0 22.000000 0 0
          8
                                                   347085 7.775000
                                                                       S
                                                                             1
          9
                  908
                        3 0 21.000000
                                                      4137 9.825000
                                                                       s
                                                                              0
                                                                       С
             313 1 1 27.000000 0 2 113503 211.500000
         10
                                                                             0
In [174]: missing_values_count = df['embarked'].isnull().sum()
print(missing_values_count)
         Θ
```

```
In [171]: df = df.drop('name', axis=1)
In [172]: df = df.drop('boat', axis=1)
In [173]: df
Out[173]:
                                                age sibsp parch
                   passenger_id pclass sex
                                                                                     ticket
                                                                                                   fare embarked survived
                         1216 3 0 29.519847 0 0
                                                                                  335432 7.733300 Q
In [176]: df = df.drop('ticket', axis=1)
In [177]: X = df.drop("survived", axis=1)
y = df["survived"]
              # Split the data into training and test sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
              # Check the shape of the training and test sets
print("Training set shape:", X_train.shape)
print("Test set shape:", X_test.shape)
              Training set shape: (680, 8)
Test set shape: (170, 8)
In [178]: missing_values_count = df['survived'].isnull().sum()
print(missing_values_count)
              0
```

0 0 0 0 1 0 0 0 1 0 0 .1323529411764706 0.2250000000000000003